Optimal income taxation is quite a different problem than optimal commodity taxation. In optimal commodity taxation the issue was which commodities to tax, and at what rates. By definition, optimal income taxation is a problem in taxing income, so we already know what we are taxing. The question then becomes what should the income tax **schedule** look like. In other words, **how progressive** should the income tax be? This problem brings out quite starkly the potential conflict between equity and efficiency.

Suppose that the government needs to raise a given amount of money from the income tax. One way of lowering the burden of the income tax on the very lowest income people is to increase the **basic exemption**. This is an amount of income which is exempt from tax. For the tax year 2013 it is \$11038. [The properties of the Canadian income tax will be discussed in more detail in section 5 of the course.] That means that you only pay federal income tax if your income is greater than \$11038. And then you pay tax only on your income above \$11038. So currently, the federal income tax payable by a person of income $11038 < y < 43561$, where y is the person's annual income, is

$$(0.15)(y - 11038)$$

That is, we have a marginal tax rate of 15 percent, and a basic credit 15 percent of 11038 (which equals \$1655.70) [If a person's income were below \$11038, she would pay no tax ; if it were above \$43561, then she would face a higher marginal rate.] A person of income 20000 would therefore have to pay federal taxes of

$$(0.15)(20000 - 11038) = 1344.30$$

If we raised the exemption to \$12000, then this person's taxes would fall to \$1200. So the person's average tax rate would fall from 6.72 % (1344.30/20000) to 6%. The results would be even stronger for people at lower incomes. A person of income \$13000 would have his taxes fall from 294.30 to 150, a reduction in his average tax rate from 2.26% to 1.15%.

Raising the basic exemption in this way would lower tax liabilities, and make the tax more progressive. But raising the exemption will lower the tax yield. Something must be done to make up the revenue shortfall. We could raise the tax collected from higher income people by raising the marginal rate they pay. However, the formula for the excess burden said that it increases with the **square** of the tax rate. Raising the marginal rate will definitely increase the excess burden of the income tax, which may already be pretty high. So there is a conflict here, and a tough choice : higher marginal rates mean a more progressive system, but also mean a greater excess burden. The question in optimal income tax theory is how to weigh these costs and benefits, and determine which factors would lead to the costs of more progressivity being higher or lower.

But the example above, and the material at the end of the section on tax incidence, 2*d* : Tax Incidence and Income Distribution, also do show that an income tax can be progressive even if it has a **constant marginal rate**. Such a tax is often referred to as a flat tax. As a rough and ready

example of why making the marginal rate higher makes the tax more progressive, assume that the average per capita income in Canada is $30,000, and that the government's revenue requirements from an income tax are $5000. If the government taxed all income above $5000 at a rate of 20 percent, this tax would collect exactly $5000 per person, on average,if the average income were $30000. On the other hand, if the government taxed all income above $10,000 at a rate of 25 percent, it would also collect the same yield, $5000 per person, if the average income were $30,000. How would different people's taxes vary with the two systems? The little table below shows the taxes paid by people (in thousands of dollars) at three income levels, $15000, $30000, and $100000, for the two different tax systems.

tax	Y	15	30	100
$t = 0.2, E = 5$		2	5	19
$t = 0.25, E = 10$		1.25	5	22.5

The higher the tax rate, the more progressive the tax.

In this example, I assumed that the average income in the country was $30,000, and more important, that that average income stayed the same whatever was the income tax system. But one of the main points of virtually every bit of the course so far has been : **taxes change people's behaviour**. People's incomes are not exogenous ; they are determined by people's career and hours worked choices, and these choices will change with the tax system. In other words, if incomes were unaffected by the nature of the income tax schedule, then there would be no excess burden to income taxation, and the equity–efficiency trade–off would disappear.

The British economist Edgeworth recognized about a century ago that the optimal income tax problem is less of a problem if the tax system did not affect the total amount of income. He asked : what would be the best income tax system, if income were truly exogenous? In particular, he wanted to know how the progressivity of the income tax would depend on people's (or the government's) attitude toward inequality. Social welfare functions often involve some aversion to inequality : transferring a dollar from a rich person to a poor person would raise social welfare, other things equal. But there can be a wide range of views on how much inequality matters. Is a dollar more to a poor person the same as $1.20 to a rich person? $2? $5? We might expect, the higher is our aversion to inequality, that is the more we think transfers to the poor are more valuable than transfers to the rich, the more progressive we would want the income tax to be.

Edgeworth showed that conclusion is **wrong** if income truly is exogenous. As long as the social welfare function values a dollar in income given to a poor person as having more social welfare than a dollar given to a rich person, the optimal income tax should involve a marginal rate of — 100 percent. It doesn't matter whether transfers to the poor are regarded as a little more valuable than transfers to the rich, or a lot. Raising the marginal income tax rate (and raising the exemption so as to keep the revenue yield constant) will transfer income from the rich to the poor. As long as the total income is unaffected, social welfare will go up.

Why does Edgeworth's result hold? A social welfare function which is inequality averse would

have the form

$$SW = W(y_1) + W(y_2) + \cdots$$

where y_i is the income of person i, and where

$$W' > 0 \quad ; \quad W'' < 0$$

It is this concavity ($W'' < 0$) of $W(\cdot)$ which makes the social welfare function risk averse : it implies that when person 1 is richer than person 2, transferring a dollar from person 1 to person 2 changes social welfare by (approximately) $W'(y_2) - W'(y_1)$. That's the effect of person 2 gaining, and person 1 losing on overall social welfare. If $W'' < 0$, then $W'(y_2) > W'(y_1)$ if $y_1 > y_2$. So a concave $W(\cdot)$ says that any transfer from a person to another person of lower income will raise SW. With a social welfare function such as this, the highest value of social welfare, consistent with the fixed total amount of income $y_1 + y_2 + \cdots$, is attained if $y_1 = y_2 = \cdots = \bar{y}$, where $\bar{y}$ is the average level of income.

How could this complete equality be achieved by the tax system? Simply have an exemption level of $\bar{y}$, and a tax rate of 100%. Then person i's after–tax income would be

$$y_i - \text{taxes} = y_i - (1)(y_i - \bar{y}) = \bar{y}$$

[where here, the income tax, unlike that in Canada, actually pays money back to people whose incomes are below the exemption level].

So if you do not think the ideal income tax system is one which confiscates everything and redistributes it equally, then probably one of two things are true : either you don't care at all about poverty or income inequality, or you think that income taxes affect economic incentives, and that these incentives are important.

The analysis becomes more difficult if we drop the (not very realistic) assumption that income is exogenous. If individuals' incomes depend on the income tax schedule that they face, then the average income has to be treated as a function of the tax system. Making the tax system more progressive will tend to lower the average income — higher marginal rates, and higher exemptions, will tend to make people work less. That means that we cannot treat the average income in the country as a constant. It varies with the income tax schedule. For instance, if the government's revenue need is $5000 per person, then if the income tax is to collect enough to cover those revenue needs, the equation

$$t[a(t, E) - E] = 5000$$

must hold, where $a(t, E)$ is the average income in the country, which depends on the marginal tax rate t and on the level of the exemption E.

Now it is not true that the marginal rate should be 100% if the social welfare function is inequality averse. The only way to transfer income from the rich to the poor is to make the tax system more progressive. But that lowers the average income. In effect, making the income tax

more progressive might lower the income of a rich person by \$2, and raise the income of a poor person by \$1. Whether that is a good idea or not depends on just how inequality averse the social welfare function is.

In other words, how progressive the income tax should be will depend on i how big an effect income taxes have on people's decisions, and ii how inequality averse the social welfare function is. This is a very different result than Edgeworth's. Even if someone was extremely averse to inequality, she might want an income tax which is much, much lower than 100 %, if high marginal tax rates had a strong negative effect on people's labour supply. For example, it might seem reasonable that hardly anyone would work (or invest) if they really did face a marginal tax rate of 100%. Why spend time and effort at work, if all the income you earn would be taxed away? If this were the case (that a 100% marginal rate would induce everyone to quit work), then the average income $a(t, E)$ would be 0 if t was 100%. So the 100% marginal rate would then leave everyone with the same after–tax income — 0. In such a case, even someone with a very strong aversion to inequality would want a tax system which does not completely equalize income : the lowest–income people would be better off with a small fraction of something than an average fraction of nothing.

So there is, in general, a serious trade–off involved in designing the income tax : higher marginal rates lead to a lower total income, but a more equitable distribution of that income.

What is that trade–off? An example of a calculation of the optimal income tax is the work of Nick Stern, done in the 1970's, and described in Rosen, Wen, and Snoddon. Stern showed was that how progressive the income tax should be will depend on our attitude towards inequality (when income is endogenous) — and also on the substitutability of leisure for other goods — and on the government's revenue requirement. In the section on "Taxation and Efficiency" the excess burden of the income tax was described as proportional to the compensated elasticity of labour supply. The bigger is that elasticity, the higher the excess burden of the income tax. That makes a progressive income tax very costly, in terms of reduced income. So the optimal income tax will not be very progressive if people's (compensated) labour supply elasticity is very high. More substitutability between leisure and goods makes this supply elasticity higher.

On the other hand, the bigger is our aversion to inequality, the higher is the marginal benefit of more progressive income tax, and the higher is the optimal income tax rate.

There is no reason to restrict attention to income tax schedules with **constant marginal tax rates**. After all, in Canada, and in many other countries, the tax schedule is not **linear**. That is, if we define a person's tax owing as a function of her income, it is not some simple function such as

$$\text{tax} = t(y - E)$$

The slope of the tax function — how much your tax increases when your income increases by \$1 — changes, from 0.15 to 0.22 to 0.26 to 0.29, if we look at the Canadian federal income tax in 2012. So we can, and do, have tax schedules which are **non–linear**, in that the marginal rate varies with income.

Is that a good idea? Some economists have looked at the optimal overall shape of the income tax schedule, without restricting that shape to be linear. That is, if we can vary the marginal rate by income class, what should be the optimal schedule? In many of these exercises it appears that it is not optimal to deviate much from a flat tax — although that conclusion is rather sensitive to the assumptions made.

The final topic in this section is the problem of **dynamic inconsistency** of the tax system, sometimes referred to as the **capital tax levy** problem. To consider this, recall that all the equity versus efficiency conflicts discussed in this section would disappear in a first–best world, that is in a world in which redistribution was possible using **lump–sum taxation**. I have stated, again and again, that the only reasonable feasible lump–sum tax is a head tax.

But what about a tax on people's wealth, or on the earnings from that wealth? If we taxed the wealth people had, then their tax liabilities would not be affected by any decisions they made. A person's wealth was determined in the past. Today's initial wealth can no more be changed by a person than can her age or her sex.

Today's wealth was acquired in the past. Nothing we do to tax its return today can alter what was done yesterday. So it is indeed true that taxation of the return to capital is lump–sum, if we only tax the return to assets that were acquired previously. That is, we could announce that there was a special tax of 40 percent on the annual return, in 2014, on all assets that were acquired prior to November 2013, to be paid by the owner of the asset as of November 15 2013. That tax, announced after November 15 2013, is lump–sum. There is nothing anyone can do to alter his or her liabilities under this tax.

The problem is that today is yesterday's tomorrow. While taxing the consequences of previous investment decisions is lump–sum, anticipation of future taxation will affect people's decisions today. To get people to invest today, the government must somehow be able to commit not to confiscate the return on that investment tomorrow. That is the dynamic inconsistency problem. In 2013, it looks best not to tax the return on investments made in 2013 ; taxing this return will have an excess burden, and will discourage current investment. In 2014, it will look efficient to tax the return on investments made in 2013, since those decisions cannot be altered in the future. So even if government today acknowledges that the future returns to investment done today should not be taxed, we all know that matters will look different tomorrow. Unless the government can somehow commit today not to do in the future what it will find efficient to do in the future, investors' rational expectations of these future taxes will depress current investment.

What does this say about the optimal tax system? Suppose that the government, in the year 2013, is designing not just the tax system for 2013 income, but the tax system for future years. Governments do very often make plans for changes in the tax system that will take place over several years in the future.

An efficient plan would be to have a very large tax on people's wealth, but only in 2013. In fact, make the tax so high that it collects much more revenue than the government needs in any one year. Perhaps even set the tax rate at 100 percent. This tax will have no excess burden, if

it is announced very late in 2013. Each person's wealth in 2013 was determined by the decisions she (and her parents and their parents) made in the past, and cannot be affected by decisions she makes now or in the future. So a tax on current wealth is lump–sum. The government can also announce that it will levy no taxes at all on wealth in the future. Maybe even no taxes on anything else, if they can collect enough revenue from the 2013 wealth tax so as to pay for their future expenditure, as well as their current expenditure. So the efficient, lump–sum tax plan is for a tax rate on wealth of 100 percent in 2013, and a tax rate on wealth of 0% in 2014, 2015, 2016 etcetera. Because wealth will not be taxed in the future, people will be willing to save so as to accumulate future wealth, since it will not be confiscated. We have a tax plan that collects a lot of money, but which has no negative efficiency consequences, by virtue of taxing only what was determined by past decisions.

But this only works if the government's plan is credible. Next year, or in 2015, the government may decide to change its plan of not taxing wealth. Why? In 2015, a tax on current wealth will be non–distortionary. Unless the government can somehow make a **binding commitment** to stick to its original (2013) plan, the plan will not be credible. Any plan for the future will be dynamically inconsistent if the government will want to change the plan. If the government cannot make such binding commitments, then it should recognize the dynamic inconsistency of its plans. Instead of formulating the tax plan which looks best from a 2013 perspective, it should formulate a plan which it knows that it will follow in the future. That is, it should set a plan in 2013, which it knows (and which everyone else knows) that it will still want to follow in 2014, and 2015, and so on. A dynamically consistent plan is one which can be expected to be followed in the future, even when the government realizes in 2015 that wealth accumulated in the previous year is now unaffected by its future decisions. The optimal dynamically consistent plan will have a bigger excess burden than the optimal plan which could be designed if the government could commit. But if investors (or potential investors) know that the government cannot commit to a plan, the dynamically consistent plan is the only feasible one.